On recurrence and transience of self-interacting random walks

Yuval Peres* Serguei Popov† Perla Sousi‡

November 25, 2012

Abstract

Let $\mu_1, \ldots, \mu_k$ be d-dimensional probability measures in $\mathbb{R}^d$ with mean 0. At each time we choose one of the measures based on the history of the process and take a step according to that measure. We give conditions for transience of such processes and also construct examples of recurrent processes of this type. In particular, in dimension 3 we give the complete picture: every walk generated by two measures is transient and there exists a recurrent walk generated by three measures.
Keywords and phrases. Transience, recurrence, Lyapunov function.
MSC 2010 *subject classifications.* Primary 60G50; Secondary 60J10.

1 Introduction

Let μ_1 and μ_2 be two zero mean measures in $\mathbb{R}^4$ with finite supports that span the whole space. On the first visit to a site the jump of the process has law μ_1 and at further visits it has law μ_2. The following question was posed in [2]: Is the resulting walk transient?

More generally, one can consider any adapted rule (i.e., a rule depending on the history of the process) for choosing between μ_1 and μ_2, and ask the same question. It turns out that the answer to this question is positive, even in $\mathbb{R}^3$, as proved in Theorem 1.2 below. Moreover, in 3 dimensions this result is sharp, in the sense that one can construct an example of a recurrent walk with three measures, as shown in Theorem 1.5.

This naturally fits into the wider context of random walks that are not Markovian, namely where the next step the walk takes also depends on the past. Recently there has been a lot of interest in random walks of this kind. A large class of such walks are the so-called vertex (or edge) reinforced random walks, where the walker chooses the next vertex to jump to with weight proportional to the number of visits to that vertex up to that time; see e.g. [1, 9, 10, 11, 13, 15]. Another class of such walks is the so-called excited random walks, when the transition probabilities depend on whether it is the first visit to a site or not, see e.g. [3, 4, 8, 14, 17].

In this paper we study transience and recurrence for walks in dimensions 3 and above that are generated by a finite collection of step distributions. We now give the precise definition of the walks we will be considering.

*Microsoft Research, Redmond, Washington, USA; peres@microsoft.com
†University of Campinas, Campinas SP, Brazil; popov@ime.unicamp.br
‡University of Cambridge, Cambridge, UK; p.sousi@statslab.cam.ac.uk

Definition 1.1. Let $\mu_1, \ldots, \mu_k$ be k probability measures in $\mathbb{R}^d$ and let $(\xi_n^j, j = 1, \ldots, k, n = 1, 2, 3, \ldots)$ be independent random variables with $\xi_n^j \sim \mu_j$ for all n and all $j = 1, \ldots, k$. Define an *adapted rule* $\ell = (\ell(i))_i$ with respect to a filtration $(\mathcal{F}_i)$ to be a process such that $\ell(i) \in \{1, \ldots, k\}$ and is $\mathcal{F}_i$ measurable for all i. We say that the walk X, with $X_0 = 0$, is generated by the measures $\mu_1, \ldots, \mu_k$ and the rule ℓ if X_i is $\mathcal{F}_i$-measurable and

$$X_{i+1} = X_i + \xi_{i+1}^{\ell(i)}.$$

We say that a measure μ in $\mathbb{R}^d$ has mean 0 if $\int_{\mathbb{R}^d} x\mu(dx) = 0$. Also we write that a measure μ has β moments, if $\mathbb{E}\left[\|Z\|^\beta\right] < \infty$, where $Z \sim \mu$. We define the covariance matrix of μ as follows: $\mathrm{Cov}(\mu) = (\mathbb{E}[Z_i Z_j])_{i,j=1}^d$.

Note that if μ is a measure in $\mathbb{R}^d$, then it has an invertible covariance matrix if and only if its support contains d linearly independent vectors of $\mathbb{R}^d$. We will call such measures d-dimensional.

We say that the walk X is transient if $\|X_n\| \to \infty$ a.s. as $n \to \infty$, and recurrent if there exists a compact set that is visited by X infinitely many times a.s. (observe that, since the walks we are considering are not necessarily Markovian, in principle one can construct examples that are neither transient nor recurrent).

In this paper we address the following two questions:

- Let $\mu_1, \ldots, \mu_k$ be mean 0 probability measures in $\mathbb{R}^d$. What are the conditions on the measures so that for every adapted rule ℓ the resulting walk is transient?
- For a given dimension d, how do we construct examples of recurrent walks generated by k d-dimensional mean 0 measures? How small can this number k be made?

In Section 1.1 we state our results concerning the first question and in Section 1.2 about the second one. Observe that Theorems 1.2 and 1.5 give a complete picture in dimension 3: any two mean 0 measures with $2 + \beta$ moments, for some $\beta > 0$, always generate a transient walk, while there is an example of a recurrent walk generated by three 3-dimensional measures of mean 0 with a suitable adapted rule.

1.1 Conditions for transience

Theorem 1.2. *Let μ_1, μ_2 be d-dimensional measures in $\mathbb{R}^d$, $d \geq 3$, with zero mean and $2 + \beta$ moments, for some $\beta > 0$. If X is a random walk generated by these measures and an arbitrary adapted rule ℓ, then X is transient.*

The following result will be used in the proof of Theorem 1.2 but is also of independent interest, since it gives a sufficient condition on the covariance matrices of the measures used in order to generate a transient random walk X for an arbitrary adapted rule ℓ.

For a matrix A we write A^T for its transpose, $\lambda_{\max}(A)$ for its maximum eigenvalue and $\mathrm{tr}(A)$ for its trace.

Theorem 1.3. *Let $\mu_1, \ldots, \mu_k$ be mean 0 measures in $\mathbb{R}^d$, $d \geq 3$, with $2 + \beta$ moments, for some $\beta > 0$. Suppose that there exists a matrix A such that for all i we have*

$$\mathrm{tr}(AM_iA^T) > 2\lambda_{\max}(AM_iA^T), \tag{1.1}$$

where M_i is the covariance matrix of the measure μ_i. If X is a random walk generated by these measures and an arbitrary adapted rule ℓ, then X is transient.

We will refer to (1.1) as the *trace condition.*

It turns out that the local central limit theorem implies the following sufficient condition for transience.

Proposition 1.4. *Let $\mu_1, \ldots, \mu_k$ be d-dimensional measures in $\mathbb{R}^d$, $d \geq 2k+1$. Then the random walk X generated by these measures and an arbitrary adapted rule ℓ is transient.*

Note there are no moment assumptions on the measures μ_i in the proposition above. We will prove Proposition 1.4 at the beginning of Section 2 and then Theorems 1.3 and 1.2 in Sections 2.1 and 2.2 respectively. In Proposition 2.6 in Section 2.3 we discuss the case when the covariance matrices are jointly diagonalizable. We present a conjectured sufficient condition for transience at the end of the paper.

1.2 Recurrence

Theorem 1.5. *There exist d mean-zero d-dimensional measures of bounded support and an adapted rule such that the corresponding walk is recurrent.*

To prove this result, it is enough to construct a particular example of a recurrent random walk in d dimensions, which is generated by d mean 0 measures that are fully supported in $\mathbb{R}^d$; we now describe this example. Let $e_0, \ldots, e_{d-1}$ be the coordinate vectors in $\mathbb{Z}^d$. We consider a random walk $(X_n, n = 0, 1, 2, \ldots)$ on $\mathbb{Z}^d, d \geq 3$, defined in the following way. Fix a parameter $\gamma > 0$, and for $x = (x_0, \ldots, x_{d-1}) \in \mathbb{Z}^d$ define $\varrho(x) = \min\{k : |x_k| = \max_{j=0,\ldots,d-1} |x_j|\}$. Then

$$X_{n+1} = X_n + \xi_{n+1},$$

where $\xi_{n+1} = \pm e_{\varrho(X_n)}$ with probabilities $\frac{\gamma}{2(\gamma+d-1)}$ and $\xi_{n+1} = \pm e_k$ for $k \neq \varrho(X_n)$ with probabilities $\frac{1}{2(\gamma+d-1)}$. In words, we choose the maximal (in absolute value) coordinate of X_n with weight γ and all the other coordinates with weight 1, and then add 1 or -1 to the chosen coordinate with equal probabilities.

We will prove Theorem 1.5 in Section 3, by showing that for each $d \geq 3$ there exists large enough γ_d such that the random walk X is recurrent for all $\gamma \geq \gamma_d$. The proof of this result relies on the explicit construction of a suitable Lyapunov function, but it is rather involved, so in Section 3 we also give simpler examples of a finite number of d-dimensional measures and adapted rules that generate a recurrent walk in d dimensions. It should be mentioned also that the fact that one can construct transient zero-mean walks in dimension 2 or recurrent ones in higher dimensions by combining sufficiently many measures, is folklore.

For instance, the following modification of [16, Remark after Lemma 3.3.27] provides an example of a transient two-dimensional walk generated by combining only two measures. Consider a random walk in $\mathbb{Z}^2$ with transition probabilities from (x, y) as follows: If $|x| \geq |y|$ then change the second coordinate (adding ± 1 equally likely) with probability 0.6 and the first coordinate with probability 0.4; if $|x| < |y|$ then change the first coordinate with probability 0.6 and the second coordinate with probability 0.4. Then it is straightforward to check that the constant 1 is an excessive measure, with invariance failing only at the origin. So, e.g. Theorem 1.9 of [12] shows that the random walk is transient.

We observe, however, that it was not previously known how to construct examples of recurrent walks in $d \geq 3$ in the most "economical" way (in the sense that the number of different measures should be minimized). Since in Theorem 1.5 only d measures are used in dimension d, this result cannot be improved in three dimensions by Theorem 1.2, and it is also quite possible that Theorem 1.5 is "optimal" in higher dimensions as well (cf. the conjecture at the end of this paper).

2 Proofs of transience

In this section we give the proofs of the results on transience. We first prove Proposition 1.4, since its proof is short and elementary.

Proof of Proposition 1.4. In order to prove this proposition, let us first give an equivalent definition of the random walk that we are considering.

For each $j = 1, \ldots, k$, let $\zeta_1^j, \zeta_2^j, \ldots$ be i.i.d. with law μ_j, also independent for different j. For an adapted rule ℓ we define for all $j \in \{1, \ldots, k\}$

$$r(j,i) = \sum_{m=1}^{i} \mathbf{1}(\ell(m) = j)$$

and then writing $\hat{r}_i = r(\ell(i), i) + 1$ we let

$$X_{i+1} = X_i + \zeta_{\hat{r}_i}^{\ell(i)}.$$

It is easy to see by induction that the process X has the same law as the process of Definition 1.1. Let $R > 0$ and for every n we define the event

$$A_n = \left\{ \exists\, i_1, \ldots, i_k \geq 0 : i_1 + \ldots + i_k = n \text{ and } \sum_{j=1}^{k} \sum_{\ell=1}^{i_j} \zeta_\ell^j \in \mathcal{B}(0,R) \right\}.$$

We now fix a choice of $i_1, \ldots, i_k$ such that $i_1 + \ldots + i_k = n$. Then by [5, Corollary/Theorem 6.2] we get for a positive constant c

$$\mathbb{P}\left(\sum_{j=1}^{k} \sum_{\ell=1}^{i_j} \zeta_\ell^j \in \mathcal{B}(0,R) \right) \leq \frac{cR^d}{n^{d/2}},$$

since there must exist some i_j which is at least n/k. It is easy to see that the total number of k-tuples $(i_1, \ldots, i_k)$ with $i_j \geq 0$ for all j and $\sum_j i_j = n$ is equal to $\binom{n+k-1}{k-1}$. Since $\binom{n+k-1}{k-1} \leq c_1 n^{k-1}$, for a positive constant c_1, we deduce that

$$\mathbb{P}(A_n) \leq c' R^d \frac{n^{k-1}}{n^{d/2}} = \frac{c' R^d}{n^{d/2-k+1}},$$

which is summable if $d \geq 2k+1$. Hence, from Borel-Cantelli we obtain that a.s. only finitely many of the events A_n happen.
Now notice that for every n we have

$$\{X_n \in \mathcal{B}(0,R)\} \subseteq A_n,$$

and hence we deduce that a.s. for all sufficiently large n, the random walk at time n will stay outside of the ball $\mathcal{B}(0,R)$. Since this is true for any $R > 0$, we get that if $d \geq 2k+1$ the random walk is transient. $\square$

2.1 Trace condition and transience

In this section we give the proof of Theorem 1.3. First we state and prove some preliminary results. The following lemma is a standard result, but we state and prove it here for the sake of completeness.

Lemma 2.1. *Let (S_t) be a random walk generated by k zero mean measures and an arbitrary adapted rule ℓ. Let $\mathcal{F}_t = \sigma(S_0, \ldots, S_t)$ be its natural filtration. Let $\alpha, r_0 > 0$ and define $\varphi(x) = \|x\|^{-\alpha} \wedge r_0^{-\alpha}$. If the process $(\varphi(S_t))$ is a super-martingale, then S is transient, in the sense that a.s.*

$$\|S_t\| \to \infty \quad \textit{as} \quad t \to \infty.$$

Proof. We first show that a.s.

$$\limsup_{t\to\infty} \|S_t\| = \infty. \tag{2.1}$$

Indeed, there exist $u \in \mathbb{S}^{d-1}$, $\varepsilon > 0$ and $h > 0$ such that for all $j \in \{1, \ldots, k\}$

$$\mathbb{P}(\langle Z_j, u\rangle > \varepsilon) \geq h,$$

where $Z_j \sim \mu_j$. This implies that for all $m, n \in \mathbb{N}$ we have

$$\mathbb{P}(\langle S_{n+m} - S_n, u\rangle > \varepsilon m \mid \mathcal{F}_n) \geq h^m.$$

Hence this shows that a.s. $\limsup_t |\langle S_t, u\rangle| \geq \varepsilon m/2$ for all m, and so (2.1) holds. Clearly, this implies that a.s.

$$\liminf_{t\to\infty} \varphi(S_t) = 0. \tag{2.2}$$

Since $(\varphi(S_t))_t$ is a positive super-martingale, the a.s. super-martingale convergence theorem gives that $\lim_{t\to\infty} \varphi(S_t)$ exists a.s. and thus from (2.2) we deduce that a.s. $\lim_{t\to\infty} \varphi(S_t) = 0$, which means that a.s. $\|S_t\| \to \infty$ as $t \to \infty$. $\square$

The following lemma shows that if the covariance matrices of the measures used to generate the walk X satisfy the trace condition (1.1), then there is a function φ such that $\varphi(X)$ is a super-martingale.

Lemma 2.2. *Let $\varphi(x) = \|x\|^{-\alpha} \wedge 1$, for $x \in \mathbb{R}^d$. Let $\mu_1, \ldots, \mu_k$ be zero mean measures in $\mathbb{R}^d$ with $2 + \beta$ moments, for some $\beta > 0$, and with covariance matrices $M_1, \ldots, M_k$ satisfying for all $i = 1, \ldots, k$*

$$\mathrm{tr}(M_i) > 2\lambda_{\max}(M_i).$$

There exists $\alpha > 0$ small enough and a constant r_0 so that if $\|x\| \geq r_0$, then for all $i = 1, \ldots, k$ if $Z_i \sim \mu_i$

$$\mathbb{E}[\varphi(x + Z_i) - \varphi(x)] \leq 0. \tag{2.3}$$

Proof. It suffices to prove (2.3) for a fixed i. Since the covariance matrix M_i is positive definite, there is an orthogonal matrix U such that UM_iU^T is diagonal with non-negative eigenvalues. The matrix UM_iU^T is the covariance matrix of the random variable UZ_i.
Since U is orthogonal, we get that for all x

$$\varphi(U(x + Z_i)) = \varphi(x + Z_i) \quad \text{and} \quad \varphi(Ux) = \varphi(x). \tag{2.4}$$

In order to prove the lemma, we will apply Taylor expansion up to second order terms to the function φ around Ux evaluated at UZ_i. We will drop the dependence on i from UZ_i and write simply Z and x instead of UZ and Ux in view of (2.4) to lighten the notation.

So, let Z have covariance matrix M which is in diagonal form and with diagonal elements $\lambda_1, \ldots, \lambda_d$. Let $\widetilde{Z} = Z\mathbf{1}(\|Z\| \le \|x\|/2)$. Note that if a.s. $\|Z\| \le B$ for a positive constant B, then $\widetilde{Z} = Z$ if $\|x\| \ge 2B$. The calculations below are a bit simpler in this case, since $\widetilde{Z}$ would have mean 0 and the same covariance matrix as Z.

If $\|x\| \ge 2$, then $\left\|x + \widetilde{Z}\right\| \ge 1$ and so $\varphi(x + \widetilde{Z}) = \left\|x + \widetilde{Z}\right\|^{-\alpha}$. In what follows we abbreviate

$$\varphi_i'(x) = \frac{\partial \varphi(x)}{\partial x_i}, \quad \varphi_{ij}''(x) = \frac{\partial^2 \varphi(x)}{\partial x_i \partial x_j}, \quad \varphi_{ijk}'''(x) = \frac{\partial^3 \varphi(x)}{\partial x_i \partial x_j \partial x_k}.$$

Applying Taylor expansion to φ up to second order terms gives for some $\eta \in (0,1)$

$$\begin{aligned}
\varphi(x + \widetilde{Z}) &= \varphi(x) + \langle \nabla\varphi(x), \widetilde{Z}\rangle + \frac{1}{2}\sum_{i,j=1}^{d} \varphi_{ij}''(x)\widetilde{Z}_i\widetilde{Z}_j + \frac{1}{3!}\sum_{i,j,k=1}^{d} \varphi_{ijk}'''(x + \eta\widetilde{Z})\widetilde{Z}_i\widetilde{Z}_j\widetilde{Z}_k \\
&= \varphi(x) + \langle \nabla\varphi(x), \widetilde{Z}\rangle + \frac{1}{2}\sum_{i,j=1}^{d} \varphi_{ij}''(x) Z_i Z_j + \frac{1}{3!}\sum_{i,j,k=1}^{d} \varphi_{ijk}'''(x + \eta\widetilde{Z})\widetilde{Z}_i\widetilde{Z}_j\widetilde{Z}_k \\
&\quad - \sum_{i,j=1}^{d} \varphi_{ij}''(x) Z_i Z_j \mathbf{1}\left(\|Z\| \ge \frac{\|x\|}{2}\right).
\end{aligned}$$

Claim 2.3. *There exist positive constants C, C_1 such that for all i, j*

$$\left|\mathbb{E}\left[\langle \nabla\varphi(x), \widetilde{Z}\rangle\right]\right| \le \frac{C}{\|x\|^{\alpha+\beta+2}} \quad \text{and} \quad \mathbb{E}[Z_i Z_j \mathbf{1}(\|Z\| \ge \|x\|/2)] \le \frac{C_1}{\|x\|^{\beta}}.$$

Proof. By Hölder's inequality we have

$$\mathbb{E}[\|Z\|\mathbf{1}(\|Z\| \ge \|x\|/2)] \le \frac{2^{\beta+1}\mathbb{E}\left[\|Z\|^{\beta+2}\right]}{\|x\|^{\beta+1}} \le \frac{K}{\|x\|^{\beta+1}}.$$

Since $\mathbb{E}[Z] = 0$, we have $\mathbb{E}\left[\widetilde{Z}\right] = \mathbb{E}\left[\widetilde{Z} - Z\right]$, and hence

$$\left\|\mathbb{E}\left[\widetilde{Z}\right]\right\| = \|\mathbb{E}[Z\mathbf{1}(\|Z\| \ge \|x\|/2)]\| \le \mathbb{E}[\|Z\|\mathbf{1}(\|Z\| \ge \|x\|/2)] \le \frac{K}{\|x\|^{\beta+1}}.$$

For the first term of the Taylor expansion we have for a positive constant C

$$\left|\mathbb{E}\left[\langle \nabla\varphi(x), \widetilde{Z}\rangle\right]\right| = \sum_{i=1}^{d} \frac{\alpha|x_i|}{\|x\|^{\alpha+2}}\left|\mathbb{E}\left[\widetilde{Z}_i\right]\right| \le \sum_{i=1}^{d} \frac{\alpha|x_i|}{\|x\|^{\alpha+2}}\left\|\mathbb{E}\left[\widetilde{Z}\right]\right\| \le \frac{\alpha d K\|x\|}{\|x\|^{\alpha+\beta+3}} = \frac{C}{\|x\|^{\alpha+\beta+2}}.$$

For all i, j we have by Hölder's inequality again

$$\mathbb{E}[Z_i Z_j \mathbf{1}(\|Z\| \ge \|x\|/2)] \le \mathbb{E}\left[\|Z\|^2 \mathbf{1}(\|Z\| \ge \|x\|/2)\right] \le \frac{C_1}{\|x\|^{\beta}},$$

thus proving the claim. □

We continue proving Lemma 2.2. For the second order terms we write

$$\mathbb{E}\Big[\widetilde{Z}_i\widetilde{Z}_j\Big] = \mathbb{E}[Z_iZ_j\mathbf{1}(\|Z\|\leq\|x\|/2)] = \mathbb{E}[Z_iZ_j] - \mathbb{E}[Z_iZ_j\mathbf{1}(\|Z\|\geq\|x\|/2)]\,,$$

and hence since for $i\neq j$ we have $\mathbb{E}[Z_iZ_j]=0$, by Claim 2.3 we get

$$\Big|\mathbb{E}\Big[\widetilde{Z}_i\widetilde{Z}_j\Big]\Big| \leq \frac{C_1}{\|x\|^\beta} \quad\text{and}\quad \left|\sum_{i=1}^d \varphi''_{ii}(x)\mathbb{E}\big[Z_i^2\mathbf{1}(\|Z\|\geq\|x\|/2)\big]\right| \leq \frac{C_2}{\|x\|^{\alpha+\beta+2}}.$$

Since for all i we have $\mathbb{E}\big[Z_i^2\big]=\lambda_i$, we obtain

$$\sum_{i=1}^d \varphi''_{ii}(x)\mathbb{E}\big[Z_i^2\big] = \sum_{i=1}^d \lambda_i\frac{-\alpha\|x\|^2+\alpha(\alpha+2)x_i^2}{\|x\|^{\alpha+4}} = \sum_{i=1}^d\frac{\alpha x_i^2(\lambda_i(\alpha+2)-\sum_{j=1}^d\lambda_j)}{\|x\|^{\alpha+4}}. \tag{2.5}$$

The rest of the second order terms can be bounded as follows:

$$\left|\sum_{i\neq j}\varphi''_{ij}(x)\mathbb{E}\Big[\widetilde{Z}_i\widetilde{Z}_j\Big]\right| = \sum_{i\neq j}\frac{\alpha(\alpha+2)|x_i||x_j|}{\|x\|^{\alpha+4}}\Big|\mathbb{E}\Big[\widetilde{Z}_i\widetilde{Z}_j\Big]\Big| \leq \sum_{i\neq j}\frac{\alpha(\alpha+2)|x_i||x_j|}{\|x\|^{\alpha+4}}\frac{C_1}{\|x\|^\beta} \leq \frac{C_3}{\|x\|^{\alpha+\beta+2}}.$$

For the remainder in the Taylor expansion we have

$$\max_{i,j,k}\Big|\varphi'''_{ijk}(x+\eta\widetilde{Z})\Big| \leq \frac{C}{\|x+\eta\widetilde{Z}\|^{\alpha+3}} \leq \frac{C_4}{\|x\|^{\alpha+3}},$$

since $\Big\|\widetilde{Z}\Big\|\leq\|x\|/2$. We want to control $\mathbb{E}[\varphi(x+Z)-\varphi(x)]$. We write

$$\mathbb{E}[\varphi(x+Z)-\varphi(x)] = \mathbb{E}\Big[\varphi(x+Z)-\varphi(x+\widetilde{Z})\Big] + \mathbb{E}\Big[\varphi(x+\widetilde{Z})-\varphi(x)\Big] \tag{2.6}$$

and by Markov's inequality since $\mathbb{E}\Big[\|Z\|^{2+\beta}\Big]<\infty$

$$\mathbb{E}\Big[\Big|\varphi(x+Z)-\varphi(x+\widetilde{Z})\Big|\Big] \leq \mathbb{P}(\|Z\|\geq\|x\|/2) \leq \frac{C_5}{\|x\|^{\beta+2}}.$$

Since $\beta>0$, if we take $0<\alpha<\beta$, then we obtain that there exists a constant $r_0>1$ so that for $\|x\|>r_0$

$$\begin{aligned}\Big|\mathbb{E}\Big[\langle\nabla\varphi(x),\widetilde{Z}\rangle\Big]\Big| + \frac{1}{2}\left|\sum_{i,j=1}^d\varphi''_{ij}(x)\mathbb{E}\left[Z_iZ_j\mathbf{1}\left(\|Z\|\geq\frac{\|x\|}{2}\right)\right]\right| + \frac{1}{3!}\sum_{i,j,k=1}^d\Big|\mathbb{E}\Big[\varphi'''_{ijk}(x+\eta\widetilde{Z})\widetilde{Z}_i\widetilde{Z}_j\widetilde{Z}_k\Big]\Big|\\ + \Big|\mathbb{E}\Big[\varphi(x+\widetilde{Z})-\varphi(x+Z)\Big]\Big| \leq \left|\frac{1}{2}\sum_{i,j=1}^d\varphi''_{ij}(x)\mathbb{E}[Z_iZ_j]\right|.\end{aligned} \tag{2.7}$$

The assumption on the trace of the matrix M gives that for α small enough (smaller than β) $\sum_{j=1}^d\lambda_j>\lambda_i(\alpha+2)$ for all i, and hence using (2.5) we get for $\|x\|\geq r_0$

$$\sum_{i=1}^d\varphi''_{ii}(x)\mathbb{E}\big[Z_i^2\big]<0.$$

This and the inequality (2.7) finishes the proof. $\square$

We now have all the required ingredients to give the proof of Theorem 1.3.

Proof of Theorem 1.3. Let $r_0 > 1$ be the constant of Lemma 2.2. Let $\widetilde{\varphi}(x) = \|x\|^{-\alpha} \wedge r_0^{-\alpha}$, for $\alpha > 0$ as in Lemma 2.2. Notice that when $\|x\| \geq r_0$, then $\widetilde{\varphi}(x) = \varphi(x) = \|x\|^{-\alpha}$. We will first show that if $Y_t = AX_t$, then

$$\mathbb{E}[\widetilde{\varphi}(Y_{t+1}) \mid \mathcal{F}_t] \leq \widetilde{\varphi}(Y_t). \tag{2.8}$$

Since $r_0 > 1$, we have $\widetilde{\varphi}(x) \leq \varphi(x)$ for all x. So we get

$$\begin{aligned}\mathbb{E}[\widetilde{\varphi}(Y_{t+1}) - \widetilde{\varphi}(Y_t) \mid \mathcal{F}_t] &= \mathbb{E}[(\widetilde{\varphi}(Y_{t+1}) - \widetilde{\varphi}(Y_t))\mathbf{1}(\|Y_t\| \geq r_0) \mid \mathcal{F}_t] \\ &\quad + \mathbb{E}[(\widetilde{\varphi}(Y_{t+1}) - \widetilde{\varphi}(Y_t))\mathbf{1}(\|Y_t\| < r_0) \mid \mathcal{F}_t] \\ &\leq \mathbb{E}[(\varphi(Y_{t+1}) - \varphi(Y_t))\mathbf{1}(\|Y_t\| \geq r_0) \mid \mathcal{F}_t],\end{aligned}$$

since $\widetilde{\varphi}(Y_t) = r_0^{-\alpha}$ if $\|Y_t\| < r_0$ and $\widetilde{\varphi}(x) \leq r_0^{-\alpha}$ for all x. Since the covariance matrices of the measures used to generate the walk Y satisfy the trace condition (1.1), Lemma 2.2 gives that

$$\mathbb{E}[(\varphi(Y_{t+1}) - \varphi(Y_t))\mathbf{1}(\|Y_t\| \geq r_0) \mid \mathcal{F}_t] \leq 0$$

and this completes the proof of (2.8). Therefore by Lemma 2.1 we get that a.s. $\|AX_t\| = \|Y_t\| \to \infty$ as $t \to \infty$. Since for all t we have $\|AX_t\| \leq \|A\| \, \|S_t\|$ and $\|A\| > 0$, we deduce that a.s.

$$\|X_t\| \to \infty \ \text{ as } t \to \infty,$$

which concludes the proof of the theorem. $\square$

2.2 Two measures in 3 dimensions

In this section we give the proof of Theorem 1.2.

Proposition 2.4. *Let M_1, M_2 be 3×3 invertible positive definite matrices. Then there exists a 3×3 matrix A such that*

$$\operatorname{tr}(AM_iA^T) > 2\lambda_{\max}(AM_iA^T) \ \forall \ i = 1, 2.$$

Proof. We prove Proposition 2.4 by constructing the matrix A of Theorem 1.3 directly.

Let μ_1, μ_2 have covariance matrices M_1 and M_2 respectively and $\xi_i \sim \mu_i$ for $i = 1, 2$. Since M_1 is positive definite, there exists an orthogonal matrix U such that UM_1U^T is diagonal, i.e.

$$UM_1U^T = \begin{pmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{pmatrix},$$

where $a, b, c > 0$ are the eigenvalues of M_1. If we now multiply the vector $U\xi_1$ by the matrix D given by

$$D = \begin{pmatrix} \frac{1}{\sqrt{a}} & 0 & 0 \\ 0 & \frac{1}{\sqrt{b}} & 0 \\ 0 & 0 & \frac{1}{\sqrt{c}} \end{pmatrix},$$

then $\mathrm{Cov}(DU\xi_1) = I$, where I stands for the 3×3 identity matrix.

So far we have applied the matrix DU to the vector ξ_1 and we have to apply the same transformation to the vector ξ_2. The vector $DU\xi_2$ will have covariance matrix $\widetilde{M}_2$. Since it is positive definite, it can be diagonalised, so there exists an orthogonal matrix V such that

$$V\widetilde{M}_2V^T = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix},$$

where $\lambda_1 \geq \lambda_2 \geq \lambda_3 > 0$ are the eigenvalues in decreasing order. Applying the same transformation to $DU\xi_1$ is not going to change its identity covariance matrix, since V is orthogonal.
The condition we want to satisfy is

$$\lambda_1 + \lambda_2 + \lambda_3 > 2\lambda_i,$$

for all $i = 1, 2, 3$. Since the eigenvalues are in decreasing order, it is clear that this inequality is always satisfied for $i = 2, 3$. Suppose that $\lambda_2 + \lambda_3 \leq \lambda_1$. Multiplying $DU\xi_2$ by the matrix

$$B = \begin{pmatrix} \frac{\sqrt{\lambda_2}}{\sqrt{\lambda_1}} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

will give us a random vector with covariance matrix

$$\begin{pmatrix} \lambda_2 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix},$$

which clearly satisfies the trace condition (1.1). Multiplying $VDU\xi_1$ by the same matrix will give us a vector with covariance matrix

$$\begin{pmatrix} \frac{\lambda_2}{\lambda_1} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

which satisfies the trace condition (1.1), since $\lambda_2 \leq \lambda_1$. □

Proof of Theorem 1.2. By projection to the first three coordinates, it is clear that it suffices to prove the theorem in 3 dimensions.
In $d = 3$, the statement of the theorem follows from Theorem 1.3 and Proposition 2.4. □

Remark 2.5. It can be seen from the proof of Proposition 2.4 that if the measures μ_1 and μ_2 are supported on any 3 dimensional subspaces of $\mathbb{R}^d$, then a walk X generated by these measures and an arbitrary adapted rule is transient.

2.3 The diagonal case

In this section we consider a particular case when for some basis of $\mathbb{R}^d$ the covariance matrices are in diagonal form and invertible. In this setting we prove that a random walk generated by $d-1$ measures and an arbitrary rule ℓ is transient.

Proposition 2.6. *Let $d \geq 4$ and $\mu_1, \ldots, \mu_{d-1}$ be mean 0 probability measures in $\mathbb{R}^d$ with $2+\beta$ moments, for some $\beta > 0$. Let $M_1, \ldots, M_{d-1}$ be their covariance matrices and suppose that $M_i M_j = M_j M_i$ for all i, j. Then there exists a $d \times d$ matrix A such that*

$$\operatorname{tr}(AM_iA^T) > 2\lambda_{\max}(AM_iA^T) \ \forall\ i \leq d-1.$$

Therefore, a random walk X generated by the measures $(\mu_i)_{i=1}^{d-1}$ and an arbitrary adapted rule ℓ is transient.

Before giving the proof of Proposition 2.6 we prove the following:

Claim 2.7. *Let $M_1, \ldots, M_k$ be $d \times d$ invertible diagonal matrices with positive entries on the diagonal. For $A \neq 0$ we define*

$$\Psi(A) = \max_{1 \leq j \leq k} \frac{\|AM_jA^T\|}{\operatorname{tr}(AM_jA^T)}. \tag{2.9}$$

Then the minimum of $\Psi(A)$ exists among all diagonal matrices A and the minimizing matrix $\widetilde{A}$ is invertible.

Proof. Since M_j is an invertible positive definite matrix, we can write $M_j = B_j B_j^T = B_j^2$, where $B_j = B_j^T$ is an invertible matrix.

Since scaling A does not change the ratio in (2.9), we may assume that $\|A\| = 1$ and restrict attention to such matrices. It is easy to see that the set $S = \{A \text{ diagonal } : \|A\| = 1\}$ is compact and the function $f_j(A) = \|AM_jA^T\|$ is continuous on S.

Let $g_j(A) = \operatorname{tr}(AM_jA^T) = \operatorname{tr}(AB_jB_j^TA^T) = |||AB_j|||^2$, where $|||C|||^2 = \sum_{i,j=1}^d c_{i,j}^2$ and we used $\operatorname{tr}(CC^T) = |||C|||^2$.

Since B_j is invertible, we have $AB_j \neq 0$ for $A \neq 0$, so g_j does not vanish on S. Thus as g_j is continuous on S, we conclude that

$$A \mapsto \max_{1 \leq j \leq k} \frac{f_j(A)}{g_j(A)}$$

is continuous on S and hence has a minimum.

Let $\widetilde{A}$ be the minimizing matrix with diagonal elements $\lambda_1, \ldots, \lambda_d \geq 0$. We will show that $\widetilde{A}$ is invertible. Suppose the contrary and assume without loss of generality that $\lambda_d = 0$.
We prove that if we replace $\lambda_d = 0$ by a small $\varepsilon > 0$, then we get a matrix $\widetilde{A}_\varepsilon$ with $\Psi(\widetilde{A}_\varepsilon) < \Psi(\widetilde{A})$.
Let the diagonal elements of M_i be $(a_j^i)_{j=1}^d$, which are all strictly positive. Then for the matrix M_i we will have for s such that $\|M_i\| = a_s^i$

$$\frac{\lambda_{\max}(\widetilde{A}M_i\widetilde{A})}{\operatorname{tr}(\widetilde{A}M_i\widetilde{A})} = \frac{\lambda_s a_s^i}{\sum_j \lambda_j a_j^i}.$$

If $\widetilde{A}_\varepsilon$ has the same elements as $\widetilde{A}$ except for the (d,d) element which is replaced by $\varepsilon > 0$ such that $\varepsilon < \frac{\lambda_i a_j^i}{a_d^i}$ for all $i = 1, \ldots, d-1$ and all $j = 1, \ldots, d-1$, then

$$\operatorname{tr}(\widetilde{A}_\varepsilon M_i \widetilde{A}_\varepsilon) = \operatorname{tr}(\widetilde{A}M_i\widetilde{A}) + \varepsilon a_d^i,$$

while $\lambda_{\max}(\widetilde{A}_\varepsilon M_i \widetilde{A}_\varepsilon) = \lambda_{\max}(\widetilde{A}M_i\widetilde{A})$.

Replacing each 0 element of $\widetilde{A}$ by a sufficiently small number gives a matrix with smaller value of Ψ, which contradicts the choice of $\widetilde{A}$. Hence this shows that $\widetilde{A}$ is invertible. $\square$

Proof of Proposition 2.6. Since $M_iM_j = M_jM_i$ for all i, j, it follows (see for instance [7, Theorem 2.5.5]) that there is one orthogonal matrix that diagonalizes all the matrices M_i. So from now on we suppose that the M_i's are diagonal.

Recall the definition of Ψ from (2.9). Let $\widetilde{A}$ be the $d \times d$ invertible matrix that minimizes Ψ among all diagonal matrices (recall Claim 2.7).

Write $\widetilde{M_i} = \widetilde{A} M_i \widetilde{A}^T$ and

$$J = \left\{ j \leq d-1 : \frac{\|\widetilde{M_j}\|}{\mathrm{tr}(\widetilde{M_j})} = \Psi(\widetilde{A}) \right\}.$$

Since $\widetilde{A}$ and M_i are diagonal invertible matrices, it follows that $\widetilde{M_i}$ is also a diagonal invertible matrix. For each $j \leq d-1$ we can find $v_j \in \mathbb{R}^d$ such that $\|v_j\| = 1$ and $\widetilde{M_j} v_j = \|\widetilde{M_j}\| v_j$. Note that since $\widetilde{M_j}$ is diagonal, it follows that v_j can be chosen to be one of the standard basis vectors of $\mathbb{R}^d$. Let $w \in \mathbb{R}^d$ have $\|w\| = 1$ and $w \perp \{v_1, \dots, v_{d-1}\}$. Then w will also be one of the standard basis vectors of $\mathbb{R}^d$.

Next, we separate two cases.
Case 1: For some $j \in J$ there is $u_j \perp v_j$ with $\|u_j\| = 1$ and $\widetilde{M_j} u_j = \|\widetilde{M_j}\| u_j$. In this case,

$$\mathrm{tr}(\widetilde{M_j}) > \langle \widetilde{M_j} v_j, v_j \rangle + \langle \widetilde{M_j} u_j, u_j \rangle = 2\|\widetilde{M_j}\|, \tag{2.10}$$

where the strict inequality follows from the fact that $\widetilde{M_j}$ is invertible. Hence in the case where $\left\|\widetilde{M_j}\right\|$ has multiplicity at least 2, we are done.

Case 2: For each $j \in J$ the leading eigenvalue $\|\widetilde{M_j}\|$ of $\widetilde{M_j}$ has multiplicity one. We will show that this case leads to a contradiction; that is we can find another matrix with smaller value of Ψ contradicting the choice of $\widetilde{A}$ as the minimizer.

Let A_ε be the $d \times d$ matrix such that $A_\varepsilon w = (1+\varepsilon) w$ and $A_\varepsilon z = z$ for all $z \perp w$. Note that A_ε will also be diagonal, since w is one of the standard basis vectors of $\mathbb{R}^d$.

Let us denote by γ_j the second largest eigenvalue of $\widetilde{M_j}$. Then the assumption of case 2 implies that for each $j \in J$ we have $\gamma_j < \|\widetilde{M_j}\|$ and $\|\widetilde{M_j} y\| \leq \gamma_j \|y\|$ for all $y \perp v_j$.

Choose $\varepsilon > 0$ such that $(1+\varepsilon)^2 \|\widetilde{M_i}\| < \mathrm{tr}(\widetilde{M_i}) \Psi(\widetilde{A})$ for all $i \notin J$ and $(1+\varepsilon)^2 \gamma_j < \|\widetilde{M_j}\|$ for all $j \in J$.

Note that since A_ε is diagonal, $A_\varepsilon^T = A_\varepsilon$ and $\widetilde{A} A_\varepsilon$ is diagonal satisfying

$$\Psi(A_\varepsilon \widetilde{A}) = \max_{1 \leq i \leq d-1} \frac{\|A_\varepsilon \widetilde{M_i} A_\varepsilon\|}{\mathrm{tr}(A_\varepsilon \widetilde{M_i} A_\varepsilon)}.$$

By completing $\{w, v_j\}$ to an orthonormal basis $\{b_m\}|_{m=1}^d$ of $\mathbb{R}^d$ we see that for all $i \leq d-1$

$$\mathrm{tr}(A_\varepsilon \widetilde{M_i} A_\varepsilon) > \mathrm{tr}(\widetilde{M_i}), \tag{2.11}$$

since $\mathrm{tr}(M) = \sum_{m=1}^d \langle M b_m, b_m \rangle$ for any matrix M and any orthonormal basis. The strict inequality follows again from the fact that the matrix $A_\varepsilon \widetilde{M_i} A_\varepsilon$ is invertible. Also

$$\|A_\varepsilon \widetilde{M_i} A_\varepsilon\| \leq \|A_\varepsilon\|^2 \|\widetilde{M_i}\| = (1+\varepsilon)^2 \|\widetilde{M_i}\|$$

and for $j \in J$ we have for all $y \perp v_j$

$$\|A_\varepsilon M_j A_\varepsilon y\| \leq (1+\varepsilon) \|M_j (A_\varepsilon y)\| \leq (1+\varepsilon) \gamma_j \|A_\varepsilon y\| \leq (1+\varepsilon)^2 \gamma_j \|y\|,$$

since $A_\varepsilon y \perp v_j$.

We conclude that $\Psi(A_\varepsilon \widetilde{A}) < \Psi(\widetilde{A})$ by considering separately in the max defining Ψ the indices $i \notin J$ and $i \in J$, and applying (2.11). This contradicts the choice of $\widetilde{A}$ as a minimizer and establishes that case 2 is impossible. $\square$

3 More measures may yield a recurrent walk

In this section we prove that the random walk described in Section 1.2 is recurrent. First we give the simpler example that was mentioned in the Introduction.

Let $\mathbb{S}^{d-1}$ be the d-dimensional unit sphere, i.e. $\mathbb{S}^{d-1} = \{x \in \mathbb{R}^d : \|x\| = 1\}$. Let $C_1, \ldots, C_k$ be caps that cover the surface of the sphere with the property that the angle between any two vectors from the origin to points on the same cap is strictly smaller than $\pi/2$. For every cap C_i, for $i = 1, \ldots, k$, we write $m(C_i)$ for the vector joining 0 to the center of the cap C_i. Then we choose $v_{i,1}, \ldots, v_{i,d-1}$ to be $d-1$ orthogonal vectors on the hyperplane orthogonal to $m(C_i)$.

For every $x \in \mathbb{R}^d$, we write $C(x)$ for the first cap in the above ordering such that the vector joining 0 and x intersects that cap and we let $C(0) = C_1$.

Theorem 3.1. *Let $0 < \varepsilon < 1$ and X be a walk in $\mathbb{R}^d$ that moves as follows. When at x it moves along the direction of $m(C(x))$ either $+1$ or -1 each with probability $1/2$ and along each of the other $d-1$ directions, i.e. along the vectors $v_{i(x),1}, \ldots, v_{i(x),d-1}$ it moves independently as follows: ± 1 with probabilities $\varepsilon/2$ and stays in place with the remaining probability. Then there exists $\varepsilon_0 > 0$ so that for all $\varepsilon \le \varepsilon_0$ the random walk X is recurrent.*

Remark 3.2. It can be shown that the ratio of the area of the unit sphere to the area of a cap as defined above with angle $\pi/2$ is equal to $2/I_{1/2}\left(\frac{d-1}{2}, \frac{1}{2}\right)$, where I is the regularized incomplete beta function. It is then elementary to obtain that the last quantity can be bounded below by $2^{d/2+1} > d$, so that in the above theorem at least $2^{d/2+1}$ measures are needed.

Proof of Theorem 3.1. We define $\varphi(x) = \log\|x\|$, for $x \in \mathbb{R}^d$. Then by Taylor expansion to second order terms we obtain for some $\eta \in (0,1)$

$$\varphi(x+Z) = \varphi(x) + \langle \nabla\varphi(x), Z\rangle + \frac{1}{2}\sum_{i,j} \frac{\partial^2\varphi(x)}{\partial x_i \partial x_j} Z_i Z_j + \frac{1}{3!}\sum_{i,j,k=1}^{d} \frac{\partial^3 \varphi(x+\eta Z)}{\partial x_i \partial x_j \partial x_k} Z_i Z_j Z_k.$$

For each i and positive constants C, C_1, since Z is bounded, we have

$$\frac{\partial\varphi}{\partial x_i} = \frac{x_i}{\|x\|^2}, \quad \frac{\partial^2\varphi}{\partial x_i^2} = \frac{\sum_{j\neq i} x_j^2 - x_i^2}{\|x\|^4} \quad \text{and} \quad \max_{i,j,k}\left|\frac{\partial^3\varphi(x+\eta Z)}{\partial x_i \partial x_j \partial x_k}\right| \le \frac{C_1}{\|x+\eta Z\|^3} \le \frac{C}{\|x\|^3}.$$

Let $u_1, \ldots, u_d$ be the vectors (basis of $\mathbb{R}^d$) as defined in the theorem. We now write both x and Z in this basis, i.e. we have that $x = \sum_{i=1}^d x_i u_i$ and $Z = \sum_{i=1}^d Z_i u_i$. Then for $i \neq j$, by independence, we get that $\mathbb{E}[Z_i Z_j] = 0$, while $\mathbb{E}\big[Z_1^2\big] = 1$ and for all $i > 1$ we have that $\mathbb{E}\big[Z_i^2\big] = \varepsilon$. Hence, putting all things together we obtain that

$$\sum_{i,j} \frac{\partial^2\varphi(x)}{\partial x_i \partial x_j}\mathbb{E}[Z_i Z_j] = \frac{(1+\varepsilon(d-3))\|x\|^2 + 2(\varepsilon-1)x_1^2}{\|x\|^4}. \tag{3.1}$$

For the first coordinate x_1 of x, when decomposed in the basis described above, we have that

$$x_1 = \|x\| \cos\theta,$$

where θ is strictly smaller than $\pi/4$, so there exists $\delta > 0$ so that $\cos\theta \geq (1+\delta)\sqrt{2}/2$. Hence, we can now bound (3.1) from above by

$$\frac{x_1^2}{\|x\|^4}\left(2(\varepsilon-1) + \frac{1}{2(1+\delta)^2}(1+\varepsilon(d-3))\right),$$

which can be made negative by choosing ε small enough. Notice that in absolute value the last expression is at least $c\|x\|^{-2}$ for a positive constant c, and hence since Z has mean 0, it follows that for $\|x\|$ large enough we have

$$\left|\frac{1}{3!}\sum_{i,j,k=1}^{d}\mathbb{E}\left[\frac{\partial^3\varphi(x+\eta Z)}{\partial x_i \partial x_j \partial x_k} Z_i Z_j Z_k\right]\right| \leq \frac{1}{2}\left|\sum_{i,j=1}^{d}\frac{\partial^2\varphi(x)}{\partial x_i \partial x_j}\mathbb{E}[Z_i Z_j]\right|.$$

Therefore we deduce that for $\|x\| \geq r_0$

$$\mathbb{E}[\varphi(x+Z) - \varphi(x)] \leq 0. \tag{3.2}$$

We now show that this implies recurrence. By the same argument used to show (2.1) we get that a.s.

$$\limsup_{t\to\infty} \|X_t\| = \infty. \tag{3.3}$$

Let $T_{r_0} = \inf\{t \geq 0 : X_t \in \mathcal{B}(0, r_0)\}$. By (3.2) we obtain that $\varphi(X_{t\wedge T_{r_0}})$ is a positive supermartingale. Hence the a.s. martingale convergence theorem gives that $\lim_{t\to\infty}\varphi(X_{t\wedge T_{r_0}}) = Y$ exists a.s. and is finite. If $T_{r_0} = \infty$ with positive probability, then since $\varphi(x) \to \infty$ as $x \to \infty$, from (3.3) we deduce that $\lim \varphi(X_{t\wedge T_{r_0}}) = \infty$ with positive probability, which is a contradiction. Therefore, $T_{r_0} < \infty$ a.s. □

We will now give the proof of Theorem 1.5.

Proof of Theorem 1.5. By [6, Theorem 2.2.1] or analogously to the last part of the proof of Theorem 3.1, to prove recurrence it is enough to find a nonnegative function f such that $f(x) \to \infty$ as $x \to \infty$, and

$$\mathbb{E}[f(X_{n+1}) - f(X_n) \mid X_n = x] \leq 0 \qquad \text{for all large enough } x. \tag{3.4}$$

Before presenting the explicit construction of such a function, let us informally explain the intuition behind this construction. First of all, a straightforward computation shows that, if Y is a simple random walk in $\mathbb{Z}^d$, then

$$\mathbb{E}[\|Y_{n+1}\| - \|Y_n\| \mid Y_n = x] = \frac{d-1}{2d}\frac{1}{\|x\|} + O(\|x\|^{-2}),$$
$$\mathbb{E}[(\|Y_{n+1}\| - \|Y_n\|)^2 \mid Y_n = x] = \frac{1}{d} + O(\|x\|^{-1}).$$

One can observe that the ratio of the drift to the second moment behaves as $\frac{d-1}{2\|x\|}$; combined with the well-known fact that the SRW is recurrent for $d = 2$ and transient for $d \geq 3$, this suggests

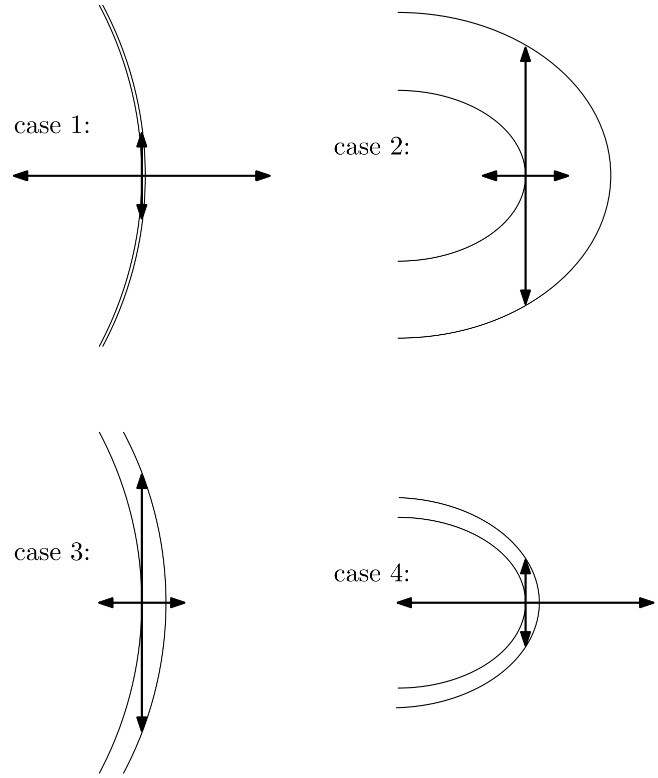

Figure 1: Looking at the level sets: how large is the drift? We have *very small* drift in case 1, *very large* drift in case 2, and *moderate* drifts in cases 3 and 4.

that, to obtain recurrence, the constant in this ratio should not be too large (in fact, at most $\frac{1}{2}$). Then, the second moment depends essentially on the dimension, and thus it is crucial to look at the drift. So, consider a (smooth in $\mathbb{R}^d \setminus \{0\}$) function $g(x) = \Theta(\|x\|)$; we shall try to figure out how the level sets of g should be so that the "drift outside" with respect to g "behaves well" (i.e., the drift multiplied by $\|x\|$ is uniformly bounded above by a not-so-large constant). For that, let us look at Figure 1: level sets of g are indicated by solid lines, vectors' sizes correspond to transition probabilities. Then, it is intuitively clear that the case of "moderate" drift corresponds to the following:

- the "preferred" direction is radial, the curvature of level lines is large, or
- the "preferred" direction is transversal and the curvature of level lines is small;

also, it is clear that "very flat" level lines always generate small drift. However, one cannot hope to make the level lines very flat everywhere, as they should go around the origin. So, the idea is to find in which places one can afford "more curved" level lines.

Observe that, for the random walk we are considering now, the preferred direction near the axes is the radial one, while in the "diagonal" regions it is in some intermediate position between transversal and radial. This indicates that the level sets of the Lyapunov function should look as depicted on Figure 2: more curved near the axes, and more flat off the axes.

We are going to use the Lyapunov function

$$f(x) = \varphi\Big(\frac{x}{\|x\|}\Big)\|x\|^\alpha,$$

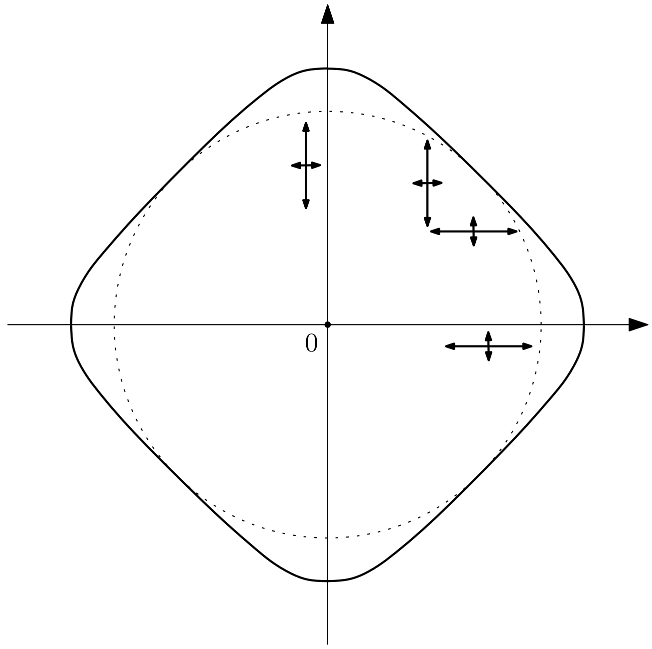

Figure 2: How the level sets of f should look like?

where α is a positive constant and $\varphi : \mathbb{S}^{d-1} \mapsto \mathbb{R}$ is a positive continuous function, symmetric in the sense that for any $(u_0, \dots, u_{d-1}) \in \mathbb{S}^{d-1}$ we have $\varphi(u_0, \dots, u_{d-1}) = \varphi(\tau_0 u_{\sigma(0)}, \dots, \tau_{d-1} u_{\sigma(d-1)})$ for any permutation σ and any $\tau \in \{-1, 1\}^d$. By the previous discussion, to have the level sets as on Figure 2, we are aiming at constructing φ with values close to 1 near the "diagonals" and less than 1 near the axes.

By symmetry, it is enough to define the function φ for $u \in \mathbb{S}^{d-1}$ such that $u_0 \geq u_{1,\dots,d-1} \geq 0$ (clearly, it then holds that $u_0 > 0$), and, again by symmetry, it is enough to prove (3.4) for all large enough $x \in \mathbb{Z}^d$ of the same kind. For such $u \in \mathbb{S}^{d-1}$ abbreviate $s_j = u_j/u_0$, $j = 1, \dots, d-1$; observe that, if $u = x/\|x\|$, then $s_j = x_j/x_0$. We are going to look for the function (for u as above) $\varphi(u) = 1 - \alpha\psi(s_1, \dots, s_{d-1})$, where ψ is a function with continuous third partial derivatives on $[0, 1]^{d-1}$ (in fact, it will become clear that the function ψ extended by means of symmetry on $[-1, 1]^d$ has continuous third derivatives on $[-1, 1]^d$; this will imply that o-s in the computations below are uniform).

Next, we proceed in the following way: we do calculations in order to figure out, which conditions the function ψ should satisfy in order to guarantee that (3.4) holds, and then try to construct a concrete example of ψ that satisfies these conditions.

First of all, a straightforward calculation shows that for any $e \in \mathbb{Z}^d$ with $\|e\| = 1$ we have

$$\|x + e\|^\alpha = \|x\|^\alpha \Big(1 + \alpha \frac{\langle x, e\rangle}{\|x\|^2} + \frac{\alpha}{2\|x\|^2} - \frac{1}{2}\alpha(2 - \alpha)\frac{\langle x, e\rangle^2}{\|x\|^2} \cdot \frac{1}{\|x\|^2} + o(\|x\|^{-2})\Big), \qquad (3.5)$$

as $x \to \infty$.

In the computations below, we will use the abbreviations

$$\psi'_j := \frac{\partial\psi(s_1, \dots, s_{d-1})}{\partial s_j}, \quad j = 1, \dots, d-1,$$
$$\psi''_{ij} := \frac{\partial^2\psi(s_1, \dots, s_{d-1})}{\partial s_i \partial s_j}, \quad i, j = 1, \dots, d-1.$$

Let us now consider $x \in \mathbb{Z}^d$. From now on we will refer to the situation when $x_0 > x_{1,\dots,d-1} \geq 0$ as the "non-boundary case" and $x_0 = x_1 = \cdots = x_m > x_{m+1} \geq \ldots \geq x_{d-1} \geq 0$ for some $m \geq 1$ as the "boundary case". Observe for the boundary case the corresponding s will be of the form $s = (1, \dots, (1)_m, s_{m+1}, \dots, s_{d-1})$; here and in the sequel we indicate the position of the symbol in a row by placing parentheses and putting a subscript. Also, in the situation when only one coordinate of the vector s changes, we use the notation of the form $\psi((\tilde{s})_j)$ for $\psi(s_1, \dots, s_{j-1}, \tilde{s}, s_{j+1}, \dots, s_{d-1})$, possibly omitting the parentheses and the subscript when the position is clear.

First we deal with the non-boundary case.
Let us consider $x \in \mathbb{Z}^d$ such that $x_0 > x_{1,\dots,d-1} \geq 0$. Again using (3.5) and observing that (recall $s_j = x_j/x_0$) $\frac{x_j}{x_0-1} = s_j(1 + x_0^{-1} + x_0^{-2} + o(\|x\|^{-2}))$ and $\frac{x_j}{x_0+1} = s_j(1 - x_0^{-1} + x_0^{-2} + o(\|x\|^{-2}))$, we write

$$
\begin{aligned}
&\mathbb{E}[f(X_{n+1}) - f(X_n) \mid X_n = x] \\
&= -(1-\alpha\psi(s))\|x\|^\alpha + \frac{\gamma}{2(\gamma+d-1)}\Bigg[\Big(1 - \alpha\psi\big(\tfrac{x_1}{x_0-1}, \dots, \tfrac{x_{d-1}}{x_0-1}\big)\Big)\|x-e_0\|^\alpha \\
&\qquad\qquad + \Big(1 - \alpha\psi\big(\tfrac{x_1}{x_0+1}, \dots, \tfrac{x_{d-1}}{x_0+1}\big)\Big)\|x+e_0\|^\alpha\Bigg] \\
&\quad + \frac{1}{2(\gamma+d-1)}\sum_{j=1}^{d-1}\Bigg[\Big(1 - \alpha\psi\big(\tfrac{x_j-1}{x_0}\big)\Big)\|x-e_j\|^\alpha + \Big(1 - \alpha\psi\big(\tfrac{x_j+1}{x_0}\big)\Big)\|x+e_j\|^\alpha\Bigg] \\
&= \|x\|^\alpha\Bigg\{\frac{\gamma}{2(\gamma+d-1)}\Bigg[\Big(1 - \alpha\psi(s) - \alpha\sum_{j=1}^{d-1}\big(\frac{s_j}{x_0} + \frac{s_j}{x_0^2}\big)\psi'_j - \frac{\alpha}{2}\sum_{i,j=1}^{d-1}\frac{s_i s_j}{x_0^2}\psi''_{ij} + o(\|x\|^{-2})\Big) \\
&\qquad\qquad \times \Big(1 - \alpha\frac{x_0}{\|x\|^2} + \frac{\alpha}{2\|x\|^2} - \frac{1}{2}\alpha(2-\alpha)\frac{x_0^2}{\|x\|^2}\cdot\frac{1}{\|x\|^2} + o(\|x\|^{-2})\Big) \\
&\qquad + \Big(1 - \alpha\psi(s) - \alpha\sum_{j=1}^{d-1}\big(-\frac{s_j}{x_0} + \frac{s_j}{x_0^2}\big)\psi'_j - \frac{\alpha}{2}\sum_{i,j=1}^{d-1}\frac{s_i s_j}{x_0^2}\psi''_{ij} + o(\|x\|^{-2})\Big) \\
&\qquad\qquad \times \Big(1 + \alpha\frac{x_0}{\|x\|^2} + \frac{\alpha}{2\|x\|^2} - \frac{1}{2}\alpha(2-\alpha)\frac{x_0^2}{\|x\|^2}\cdot\frac{1}{\|x\|^2} + o(\|x\|^{-2})\Big) \\
&\qquad - 2(1-\alpha\psi(s))\Bigg] \\
&\quad + \frac{1}{2(\gamma+d-1)}\sum_{j=1}^{d-1}\Bigg[-2(1-\alpha\psi(s)) + \Big(1 - \alpha\psi(s) + \alpha x_0^{-1}\psi'_j - \frac{\alpha}{2x_0^2}\psi''_{jj}\Big) \\
&\qquad\qquad \times \Big(1 - \alpha\frac{x_j}{\|x\|^2} + \frac{\alpha}{2\|x\|^2} - \frac{1}{2}\alpha(2-\alpha)\frac{x_j^2}{\|x\|^2}\cdot\frac{1}{\|x\|^2} + o(\|x\|^{-2})\Big) \\
&\qquad + \Big(1 - \alpha\psi(s) - \alpha x_0^{-1}\psi'_j - \frac{\alpha}{2x_0^2}\psi''_{jj}\Big) \\
&\qquad\qquad \times \Big(1 + \alpha\frac{x_j}{\|x\|^2} + \frac{\alpha}{2\|x\|^2} - \frac{1}{2}\alpha(2-\alpha)\frac{x_j^2}{\|x\|^2}\cdot\frac{1}{\|x\|^2} + o(\|x\|^{-2})\Big)\Bigg]\Bigg\} \\
&= \alpha\|x\|^\alpha\Bigg\{\frac{\gamma}{\gamma+d-1}\Bigg[\frac{1-\alpha\psi(s)}{2\|x\|^2} - \frac{(2-\alpha)(1-\alpha\psi(s))}{2\|x\|^2}\cdot\frac{x_0^2}{\|x\|^2} - \sum_{j=1}^{d-1}\Big(\frac{s_j}{x_0^2} - \frac{\alpha s_j}{\|x\|^2}\Big)\psi'_j
\end{aligned}
$$

$$
\begin{aligned}
&\qquad\qquad\qquad\qquad - \frac{1}{2}\sum_{i,j=1}^{d-1} \frac{s_i s_j}{x_0^2}\psi''_{ij} + o(\|x\|^{-2})\Big]\\
&\quad + \frac{1}{\gamma+d-1}\Big[\frac{(d-1)(1-\alpha\psi(s))}{2\|x\|^2} + \frac{(2-\alpha)(1-\alpha\psi(s))}{2\|x\|^2}\cdot\frac{x_0^2}{\|x\|^2} - \frac{(2-\alpha)(1-\alpha\psi(s))}{2\|x\|^2}\\
&\qquad\qquad\qquad\qquad - \sum_{j=1}^{d-1}\frac{\alpha s_j}{\|x\|^2}\psi'_j - \frac{1}{2}\sum_{j=1}^{d-1}\frac{1}{x_0^2}\psi''_{jj} + o(\|x\|^{-2})\Big]\Big\}\\
&= -\alpha\|x\|^{\alpha-2}\Phi(x,\psi) + \alpha\|x\|^{\alpha-2}\big(\gamma^{-1}\Phi_1(x,\psi,\gamma,\alpha) + \alpha\Phi_2(x,\psi,\gamma,\alpha)\big), \qquad (3.6)
\end{aligned}
$$

where Φ_1 and Φ_2 are uniformly bounded for large enough x, and

$$
\Phi(x,\psi) = \frac{x_0^2}{\|x\|^2} - \frac{1}{2} + \frac{\|x\|^2}{x_0^2}\Big(\sum_{j=1}^{d-1} s_j\psi'_j + \frac{1}{2}\sum_{i,j=1}^{d-1} s_i s_j \psi''_{ij}\Big). \qquad (3.7)
$$

The idea is then to prove that, with a suitable choice for ψ, the quantity $\Phi(x,\psi)$ will be uniformly positive for all large enough x, and then the second term in the right-hand side of (3.6) can be controlled by choosing large γ and small α. This will make (3.6) negative for all large x.

Now, in order to obtain a simplified form for (3.7), we pass to the (hyper)spherical coordinates:

$$
\begin{aligned}
s_1 &= r\cos\theta_1,\\
s_2 &= r\sin\theta_1\cos\theta_2,\\
&\dots\\
s_{d-2} &= r\sin\theta_1\dots\sin\theta_{d-3}\cos\theta_{d-2},\\
s_{d-1} &= r\sin\theta_1\dots\sin\theta_{d-3}\sin\theta_{d-2}.
\end{aligned}
$$

Since $\frac{\|x\|^2}{x_0^2} = 1 + r^2$, and (abbreviating $\psi'_r = \frac{\partial\psi}{\partial r}$ and $\psi''_{rr} = \frac{\partial^2\psi}{\partial r^2}$)

$$
\psi'_r = \frac{1}{r}\sum_{j=1}^{d-1} s_j\psi'_j, \qquad \psi''_{rr} = \frac{1}{r^2}\sum_{i,j=1}^{d-1} s_i s_j\psi''_{ij},
$$

we have

$$
\begin{aligned}
\Phi(x,\psi) &= \frac{1}{1+r^2} - \frac{1}{2} + (1+r^2)\Big(r\psi'_r + \frac{r^2}{2}\psi''_{rr}\Big)\\
&= \frac{1+r^2}{2}\Big(\frac{1-r^2}{(1+r^2)^2} + \big(r^2\psi'_r\big)'_r\Big). \qquad (3.8)
\end{aligned}
$$

Now, we define the function ψ (it will depend on r only, not on $\theta_1,\dots,\theta_{d-2}$) in the following way. First, clearly, we need to define $\psi(r)$ for $r \in [0, \sqrt{d-1}]$. Then, observe that

$$
\int_0^{\sqrt{d-1}} \frac{1-r^2}{(1+r^2)^2}\,dr = \frac{\sqrt{d-1}}{d} > 0, \qquad (3.9)
$$

so, for a suitable (small enough) ε_0 we can construct a smooth function h with the following properties (on the Cartesian plane with coordinates (r,y), think of going from the origin along $y = \frac{r^2}{4\varepsilon_0^2}$ until it intersects with $y = \frac{1-r^2}{(1+r^2)^2}$ and then modify a little bit the curve around the intersection point to make it smooth, see Figure 3):

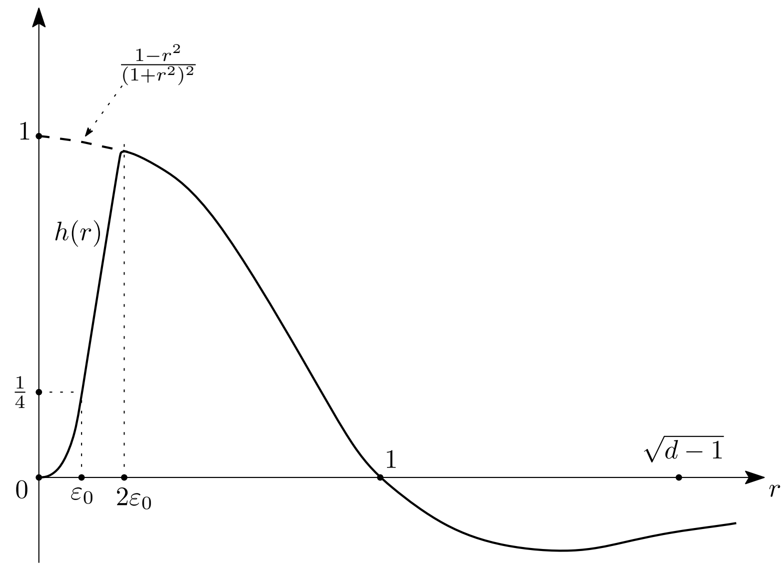

Figure 3: On the construction of h

(i) $0 \le h(r) \le \frac{1-r^2}{(1+r^2)^2}$ for all $r < 2\varepsilon_0$ and $h(r) = \frac{1-r^2}{(1+r^2)^2}$ for $r \ge 2\varepsilon_0$;

(ii) $h(0) = 0$ and $h(r) \sim \frac{r^2}{4\varepsilon_0^2}$ as $r \to 0$;

(ii) $\frac{1-r^2}{(1+r^2)^2} - h(r) > \frac{1}{2}$ for $r \le \varepsilon_0$;

(iv) $b := \int_0^{\sqrt{d-1}} h(r)\,dr > 0$ (by (3.9) it holds in fact that $b \in (0,1)$);

(v) $\int_0^r h(u)\,du > \frac{br^3}{3(d-1)^{3/2}}$ for all $r \in (0, \sqrt{d-1}]$.

Denote $H(r) = \int_0^r h(u)\,du$, so that we have $H(\sqrt{d-1}) = b$. Then, define for $r \in [0, \sqrt{d-1}]$

$$\psi(r) = \int_r^{\sqrt{d-1}} \Big(\frac{H(v)}{v^2} - \frac{bv}{3(d-1)^{3/2}}\Big)\,dv. \tag{3.10}$$

For the function ψ defined in this way, we have $r^2\psi'(r) = \frac{br^3}{3(d-1)^{3/2}} - H(r)$, so $h(r) + (r^2\psi'(r))' = b(d-1)^{-3/2}r^2$. By construction, it then holds that

$$\inf_{r\in[0,\sqrt{d-1}]} \Big(\frac{1-r^2}{(1+r^2)^2} + \big(r^2\psi'(r)\big)'\Big) \ge b(d-1)^{-3/2}\varepsilon_0^2 \wedge \frac{1}{2}, \tag{3.11}$$

and this (recall (3.7) and (3.8)) shows that, if γ is large enough and α is small enough then the right-hand side of (3.6) is negative for all large enough $x \in \mathbb{Z}^d$.

To complete the proof of the theorem, it remains to deal with the boundary case.
Let $x_0 = x_1 = \cdots = x_m > x_{m+1} \ge \ldots \ge x_{d-1} \ge 0$ for some $m \ge 1$. Using (3.5) (up to the term of order $\|x\|^{-1}$ in the parentheses), using the fact that φ is invariant under permutations and observing that x_0 and $\|x\|$ are of the same order, we have

$$
\begin{aligned}
&\mathbb{E}[f(X_{n+1}) - f(X_n) \mid X_n = x] \\
&= -(1-\alpha\psi(s))\|x\|^\alpha + \frac{\gamma+m}{2(\gamma+d-1)}\Big[\Big(1-\alpha\psi\big((\tfrac{x_0-1}{x_0})_m\big)\Big)\|x-e_0\|^\alpha \\
&\qquad + \Big(1-\alpha\psi\big(\tfrac{x_0}{x_0+1},\dots,(\tfrac{x_0}{x_0+1})_m,\tfrac{x_{m+1}}{x_0+1}\dots,\tfrac{x_{d-1}}{x_0+1}\big)\Big)\|x+e_0\|^\alpha\Big] \\
&\quad + \frac{1}{2(\gamma+d-1)}\sum_{j=m+1}^{d-1}\Big[\Big(1-\alpha\psi\big(\tfrac{x_j-1}{x_0}\big)\Big)\|x-e_j\|^\alpha + \Big(1-\alpha\psi\big(\tfrac{x_j+1}{x_0}\big)\Big)\|x+e_j\|^\alpha\Big] \\
&= \|x\|^\alpha\Big\{\frac{\gamma+m}{2(\gamma+d-1)}\Big[\Big(1-\alpha\Big(\psi(s)-\frac{\psi'_m}{x_0}+o(\|x\|^{-1})\Big)\Big)\Big(1-\alpha\frac{x_0}{\|x\|^2}+o(\|x\|^{-1})\Big) \\
&\qquad + \Big(1-\alpha\Big(\psi(s)-\sum_{k=1}^{m}\frac{\psi'_k}{x_0}-\sum_{k=m+1}^{d-1}\frac{s_k\psi'_k}{x_0}+o(\|x\|^{-1})\Big)\Big)\Big(1+\alpha\frac{x_0}{\|x\|^2}+o(\|x\|^{-1})\Big) \\
&\qquad - 2(1-\alpha\psi(s))\Big] \\
&\quad + \frac{1}{2(\gamma+d-1)}\sum_{j=m+1}^{d-1}\Big[\Big(1-\alpha\Big(\psi(s)-\frac{\psi'_j}{x_0}+o(\|x\|^{-1})\Big)\Big)\Big(1-\alpha\frac{x_j}{\|x\|^2}+o(\|x\|^{-1})\Big) \\
&\qquad \Big(1-\alpha\Big(\psi(s)+\frac{\psi'_j}{x_0}+o(\|x\|^{-1})\Big)\Big)\Big(1+\alpha\frac{x_j}{\|x\|^2}+o(\|x\|^{-1})\Big) - 2(1-\alpha\psi(s))\Big]\Big\} \\
&= \alpha\|x\|^\alpha\frac{\gamma+m}{2(\gamma+d-1)}\Big[\frac{1}{x_0}\Big(\sum_{k=1}^{m-1}\psi'_k + 2\psi'_m + \sum_{k=m+1}^{d-1}s_k\psi'_k\Big) + o(\|x\|^{-1})\Big] \qquad (3.12)
\end{aligned}
$$

(observe that in the above calculation all the terms of order $\|x\|^{\alpha-1}$ that correspond to the choice of coordinates $m+1,\dots,d-1$ of x, cancel).

Now simply note that by the property (v), we have $\psi'(r) < 0$ for all $r \in (0,\sqrt{d-1}]$. Observe also that for some positive constant δ_0 it holds that $\psi'(r) \le -\delta_0$ for all $r \in [1,\sqrt{d-1}]$. Then (recall that in the boundary case $s_1 = 1$ and $s_j \ge 0$ for all $j = 2,\dots,d-1$) we have

$$
\psi'_j = \frac{s_j}{r}\psi'_r \le 0 \text{ for all } j = 1,\dots,d-1 \quad \text{ and } \quad \psi'_1(s) \le -\frac{\delta_0}{\sqrt{d-1}}.
$$

This implies that the right-hand side of (3.12) is negative for all large enough $x \in \mathbb{Z}^d$ and thus concludes the proof of Theorem 1.5. □

A conjecture

We end this paper with an open question:

Conjecture 3.3. *Let $\mu_1,\dots,\mu_{d-1}$ be d-dimensional measures in $\mathbb{R}^d$, $d \ge 4$, with 0 mean and $2+\beta$ moments, for some $\beta > 0$, and ℓ an arbitrary adapted rule. Then the walk X generated by these measures and the rule ℓ is transient.*

To answer this question, by Theorem 1.3 it suffices to prove the existence of a matrix A satisfying the trace condition (1.1). So far, we were able to prove it in the case when the $d-1$ covariance matrices are jointly diagonalizable.

Acknowledgements

We thank Itai Benjamini for asking the question that led to this work and the organizers of the XV Brazilian Probability School where this collaboration was initiated. We also thank Ronen Eldan, Gady Kozma and Miklos Racz for helpful discussions. The last two authors thank Microsoft Research, Redmond, and MSRI, Berkeley, where this work was completed, for their hospitality. The work of Serguei Popov was partially supported by CNPq (300328/2005–2) and FAPESP (2009/52379–8).

References

[1] Michel Benaïm. Vertex-reinforced random walks and a conjecture of Pemantle. *Ann. Probab.*, 25(1):361–392, 1997.

[2] Itai Benjamini, Gady Kozma, and Bruno Schapira. A balanced excited random walk. *C. R. Math. Acad. Sci. Paris*, 349(7-8):459–462, 2011.

[3] Itai Benjamini and David B. Wilson. Excited random walk. *Electron. Comm. Probab.*, 8:86–92 (electronic), 2003.

[4] Jean Bérard and Alejandro Ramírez. Central limit theorem for the excited random walk in dimension $D \geq 2$. *Electron. Comm. Probab.*, 12:303–314 (electronic), 2007.

[5] C. G. Esseen. On the concentration function of a sum of independent random variables. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 9:290–308, 1968.

[6] G. Fayolle, V. A. Malyshev, and M. V. Menshikov. *Topics in the constructive theory of countable Markov chains.* Cambridge University Press, Cambridge, 1995.

[7] Roger A. Horn and Charles R. Johnson. *Matrix analysis.* Cambridge University Press, Cambridge, 1990. Corrected reprint of the 1985 original.

[8] M. Menshikov, S. Popov, A. Ramirez, and M. Vachkovskaia. On a general many-dimensional excited random walk. *Ann. Probab.*, 40(5):2106–2130, 2012.

[9] Franz Merkl and Silke W. W. Rolles. Recurrence of edge-reinforced random walk on a two-dimensional graph. *Ann. Probab.*, 37(5):1679–1714, 2009.

[10] Robin Pemantle and Stanislav Volkov. Vertex-reinforced random walk on $\mathbf{Z}$ has finite range. *Ann. Probab.*, 27(3):1368–1388, 1999.

[11] Olivier Raimond and Bruno Schapira. On some generalized reinforced random walk on integers. *Electron. J. Probab.*, 14:no. 60, 1770–1789, 2009.

[12] D. Revuz. *Markov chains.* North-Holland Publishing Co., Amsterdam, second edition, 1984.

[13] Pierre Tarrès. Vertex-reinforced random walk on $\mathbb{Z}$ eventually gets stuck on five points. *Ann. Probab.*, 32(3B):2650–2701, 2004.

[14] Remco van der Hofstad and Mark Holmes. Monotonicity for excited random walk in high dimensions. *Probab. Theory Related Fields*, 147(1-2):333–348, 2010.

[15] Stanislav Volkov. Vertex-reinforced random walk on arbitrary graphs. *Ann. Probab.*, 29(1):66–91, 2001.

[16] Ofer Zeitouni. Random walks in random environment. In *Lectures on probability theory and statistics*, volume 1837 of *Lecture Notes in Math.*, pages 189–312. Springer, Berlin, 2004.

[17] Martin P. W. Zerner. Recurrence and transience of excited random walks on $\mathbb{Z}^d$ and strips. *Electron. Comm. Probab.*, 11:118–128 (electronic), 2006.